心理科学进展 2021, Vol. 29, No. 8, 1331–1344
Advances in Psychological Science
https://doi.org/10.3724/SP.J.1042.2021.01331

• 主编特邀(Editor-In-Chief Invited) •

编者按：新世纪头 20 年，国内心理统计方法研究取得长足进步，为使读者了解国内心理统计方法研究对学科发展的大致影响、贡献和局限，为未来的方法研究提供资料和借鉴，本刊特邀温忠麟教授研究团队对 2001~2020 年发表在国内心理学专业期刊上的统计方法论文作了分类和回顾，并撰著成篇。该研究工作量不小，是一项很有意义的工作。对方法研究有兴趣的读者，一篇在手，可以总览全局，了解近 20 年国内的研究概况和发展趋势；对于应用工作者，该文可以充当查找方法文献的指南，便于根据自己想做的分析查找相应的方法论文。

新世纪 20 年国内心理统计方法研究回顾*

温忠麟[1] 方 杰[2] 沈嘉琦[1] 谭倚天[1] 李定欣[1] 马益铭[1]

([1]华南师范大学心理学院/心理应用研究中心，广州 510631)
([2]广东财经大学人文与传播学院，广州 510320)

摘 要 新世纪头 20 年，国内心理学 11 本专业期刊一共发表了 213 篇统计方法研究论文。研究范围主要包括以下 10 类(按论文篇数排序)：结构方程模型、测验信度、中介效应、效应量与检验力、纵向研究、调节效应、探索性因子分析、潜在类别模型、共同方法偏差和多层线性模型。对各类做了简单的回顾与梳理。结果发现，国内心理统计方法研究的广度和深度都不断增加，研究热点在相互融合中共同发展；但综述类论文比例较大，原创性研究论文比例有待提高，研究力量也有待加强。

关键词 心理统计，文献计量，内容分析
分类号 B841

1 前言

统计分析在心理学研究中发挥着重要作用。崔丽霞和郑日昌(2001)对 1981~1998 年发表在《心理学报》、《心理科学》和《心理发展与教育》上的论文所使用的研究方法进行了回顾：经常被使用的定量分析方法是显著性检验，包括正态检验、t 检验、χ^2 检验和 F 检验。焦璨、黄泽娟等(2010)对 1998~2008 年发表在《心理学报》和《心理科学》上使用了统计方法的论文进行了文献分析：使用最多的依次是方差分析，t 检验和相关分析；多元统计使用最多的是因子分析。此外，还有研究者对特定统计方法在心理学研究中的使用情况进行文献分析(例如，范津砚 等，2003；赵必华，顾海根，2010；方杰 等，2013；唐文清，方杰 等，2014)。上述研究都是对统计方法在心理学研究中的应用情况进行文献分析，未见有对心理统计方法研究的文献分析。本文对 2001~2020 这 20 年间发表在国内心理学 11 本专业期刊(见表 1)上的心理统计方法研究论文进行回顾和述评，将有助读者总览过去 20 年国内心理统计方法研究的总体情况和发展趋势，认识存在的问题和不足，对于心理统计学的发展很有意义。本文不涉及国内研究者发表在非心理学期刊上的论文或者海外期刊上的论文。为了节省篇幅，本文没有使用英文期刊文献，有需要的读者可以通过本文参考文献查找到相关英文期刊文献。

2 研究方法

2.1 论文来源及入选标准

论文来自国内 11 本心理学专业期刊(见表 1)上的心理统计方法论文。以中国知网(https://www.

收稿日期：2021-03-11
* 国家自然科学基金项目(31771245)、国家社会科学基金项目(17BTJ035)资助。
通信作者：温忠麟, E-mail: wenzl@scnu.edu.cn

cnki.net/)全文数据库为数据源，出版年限设为2001年1月1日~2020年12月31日，检索得到的心理统计方法类论文为研究对象。心理测量的方法研究(包括概化理论、项目反应理论和认知诊断理论)和元分析方法研究不在检索范围内。最终获得符合条件的论文213篇，发表期刊和年份的具体分布见表1。其中，发文最多的期刊前4位依次是：《心理科学》63篇，《心理科学进展》55篇，《心理学报》35篇和《心理学探新》34篇。从出版时间上看，2001~2005年35篇，2006~2010年40篇，2011~2015年85篇，2016~2020年53篇。后10年的发文数(138篇)远多于前10年(75篇)，仅2011~2015这5年间发表的就超过了之前10年的总和。但是，相比2011~2015年，2016~2020年发文数有所下降。

2.2 编码

根据内容所涉及的最主要统计方法对文章进行编码和归类。有5篇以上文章讨论同一方法的归为一类(每类可视为一个研究热点)，否则归为“其他”类。有关结构方程模型(structural equation model)39篇、测验信度(reliability) 27篇、中介效应(mediation effect) 25篇(其中包含中介与调节的整合分析5篇)、效应量(effect size)与检验力(testing power)23篇、纵向(追踪)研究(longitudinal study)18篇、调节效应(moderation effect)15篇(其中包含中介与调节的整合分析5篇)、探索性因子分析(exploratory factor analysis)11篇、潜在类别分析(latent class analysis)10篇、共同方法偏差(common method bias) 8篇、多层线性模型(hierarchical linear model)8篇、其他34篇。

3 结果

新世纪头20年，国内心理统计方法研究按论文篇数形成了10个研究热点，下面做一个简单回顾。

3.1 结构方程模型

心理统计方法研究中，有关结构方程模型(验证性因子模型是其特例)的论文最多，可以分为5个方面：1)模型拟合评价和模型选择; 2)测量指标的合并(即题目打包); 3)模型估计; 4)测量不变性; 5)结构方程建模的拓展。

温忠麟、侯杰泰和马什赫伯特(2004)指出，一个好的拟合指数应当具有3个特征，即与样本容量 N 无关、惩罚复杂的模型和对误设模型敏感。通过比较，他们建议使用NNFI (也称为TLI)和CFI (临界值为0.9), RMSEA (临界值为0.08)。郭庆科等(2007, 2008)比较了不同的拟合指标在验证性因子分析的模型拟合中的表现，还利用模拟方法得出新的拟合指数临界值标准，但温忠麟和侯杰泰(2008)用简单的 Z 检验说明检验的临界值标准是不能通过模拟研究来确定的。温涵和梁韵斯(2015)建议报告RMSEA、NNFI和SRMR，并指

表1 2001~2020年国内心理统计方法论文的频数分布(按发文篇数排序)

刊名	年份																				合计
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	
心理科学	1	3	0	1	11	1	4	3	2	4	5	3	6	3	4	2	3	4	1	2	63
心理科学进展	0	2	3	2	0	0	3	0	2	5	7	4	2	6	2	2	3	3	4	5	55
心理学报	1	1	1	2	2	1	0	4	1	0	5	4	2	3	1	2	4	0	1	0	35
心理学探新	0	1	1	0	0	0	2	0	0	3	3	2	4	5	5	1	2	2	3	0	34
心理技术与应用	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	3	0	6
中国临床心理学杂志	2	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	5
心理发展与教育	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	2	0	0	0	5
中国心理卫生杂志	0	0	0	0	1	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	4
心理与行为研究	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0	0	1	0	0	0	4
应用心理学	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
心理研究	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
合计	4	7	5	5	14	6	9	8	5	12	21	14	16	21	13	8	15	11	12	7	213

出当 NNFI 准则(临界值 0.9)接受了模型, 则无需报告 CFI。王阳等(2020)建议将等效性检验用于结构方程模型评价。该方法将推断统计元素融入了 CFI 和 RMSEA, 可以用于单个模型拟合评价和测量不变性检验。还有, 柳恒超等(2007)讨论了嵌套模型情况下的模型选择(如似然比检验法)和非嵌套模型情况下的模型选择问题(如 AIC 指标法)。

卞冉等(2007)从题目组合的基本逻辑、优缺点以及组合方法等方面对题目组合的研究进行了概括。吴艳和温忠麟(2011a)系统述评了因子法、相关法、对称法、随机法等 6 种打包策略, 给出了打包法的前提条件, 即单维和同质; 建议对于多维量表, 在子量表内打包。

测量不变性(等价性)是指观测指标和潜变量(因子)之间的关系在相比较的各个组间等同。白新文和陈毅文(2004)介绍了测量等价性的概念及其研究历史、在结构方程模型中测量等价性所要满足的 5 个条件和模型优劣判定的拟合度指数。刘军和吴维库(2005)介绍了测量不变性的各种形式, 并演示了应用结构方程模型评估测量不变性的过程。郑显亮等(2011)通过一个实例演示了如何进行二阶因子模型的测量等价性检验。李冲和刘红云(2011)考察了样本量、组间样本量对比情况、阈值差异程度、量表长度等因素对等级数据的测量不变性检验效果的影响。刘红云和李冲等(2012)比较了单维测验情境下基于分类数据的验证性因子模型和基于项目反应理论的测量不变性检验方法, 以及多维测验情境下基于分类数据的验证性因子模型和基于多维项目反应理论模型的卡方检验方法的差异。温聪聪等(2019)讨论了当截距恒定假设因其过于苛刻而无法满足时, 可以使用对齐法, 通过检验设定的模型参数的近似恒定性而进行跨群组因子均值比较。

梁莘娅和杨艳云(2016)比较了稳健极大似然估计和贝叶斯估计在结构假设和分布假设不满足时的验证性因子分析中的表现, 结果发现当数据呈非正态分布时, 贝叶斯估计法更优。但当样本量不足且数据呈正态分布时, 稳健极大似然估计法更好。因子分析假设观测变量是连续和正态分布的, 观测变量和因子是线性相关的。如果假设违背, 还可以采用项目反应理论进行题目因子分析(吴瑞林, 涂冬波, 2013; 俞宗火 等, 2006)。刘红云等人比较了分类数据的验证性因子分析和项目反应理论在单维和多维题目因子分析的参数估计中的表现(刘红云, 李美娟 等, 2012; 刘红云, 骆方 等, 2012)。

结构方程建模的拓展主要体现在验证性因子分析模型的拓展, 包括探索性结构方程模型, 因子混合模型、高阶因子模型和双因子模型。麦玉娇和温忠麟(2013)介绍了探索性结构方程模型及其主要功能和优势。探索性结构方程模型的测量模型类似于探索性因子分析(允许有跨因子负荷), 同时保留了验证性因子分析的功能。陈宇帅等(2015)介绍了因子混合模型的数学模型、基本原理, 归纳了因子混合模型的主要优势以及实际应用。因子混合模型是验证性因子模型和潜在类别模型的整合, 它能描述不同潜在类别的因子模型。

双因子模型是一种既有全局因子又可以有多个局部因子的模型(温忠麟, 叶宝娟, 2011)。顾红磊等(2014)讨论了双因子模型在信度研究、平衡量表、探索性因子分析等方面的应用。双因子模型的研究主要分为三个方面。第一、双因子模型和高阶因子模型的比较(徐霜雪 等, 2017; 温忠麟 等, 2019), 研究什么情况下使用双因子模型进行预测比较好。第二、双因子模型和随机截距因子分析模型的比较(韦嘉 等, 2016)。第三、利用双因子模型给出报告测验总分和维度分的策略, 多维量表何时可以合成总分, 何时只能报告各维度合成分数(刘玥, 刘红云, 2017; 顾红磊, 温忠麟, 2017)。

结构方程模型的其它拓展方向如下。王晓丽等(2011)指出, 根据指标和潜变量之间的关系, 结构方程的测量模型可分为形成性测量模型(指标影响潜变量)和反映性测量模型(潜变量影响指标)。任学柱等(2017)使用结构方程建模预设路径模型, 此模型可用于认知心理学的实验数据分析。还有研究者使用结构方程建模行动者-对象互依性模型, 用于成对数据(如夫妻数据、亲子数据)分析(李育辉, 黄飞, 2010)和成对模式(主体模式、客体模式、对偶模式与对比模式)检验(刘畅, 伍新春, 2017)。王孟成、邓俏文和毕向阳(2017)以一个验证性因子分析为例, 介绍了贝叶斯结构方程建模的分析过程。张沥今等(2019)介绍了几类常用的贝叶斯结构方程模型及其应用现状。

此外, 还有其他讨论结构方程模型的论文(方平 等, 2001; 方平 等, 2002; 李旭 等, 2001; 赵

必华, 顾海根, 2010; 郑文智, 吴文毅, 2014)。至于中介模型、调节模型、纵向模型和多层线性模型中的结构方程建模研究参见本文的相应部分。

3.2 信度分析

信度分析类的研究主要包括 3 个方面：α 系数与内部一致性、同质性的辨析; 信度的置信区间研究; 多层和纵向数据的信度研究。

传统教科书将 α 系数称为内部一致性信度或者同质性系数, 这其实是个误解。孟庆茂和刘红云(2002)讨论了 α 系数作为测验同质性信度的局限性(也见刘红云, 2008)。温忠麟和叶宝娟(2011)明确指出 α 系数既不能用来衡量测验的内部一致性, 也不能用来衡量测验的同质性; 在很一般的条件下, α 系数和同质性系数都不超过内部一致性信度(又称为合成信度), 后者不超过测验信度。

屠金路等(2005)展示了如何使用 Bootstrap 法进行单维测验合成信度的区间估计。叶宝娟和温忠麟(2011)比较了 Bootstrap 法、Delta 法和直接用统计软件输出的标准误在计算单维测验合成信度的置信区间中的表现, 推荐用 *Mplus* 软件估计 Delta 法的单维测验合成信度的置信区间。杨强等(2014b)给出了 SPSS 软件窗口操作计算单维测验的合成信度的方法。叶宝娟和杨强(2015)探讨了如何用 Delta 法计算一般的单维测验合成信度及其置信区间。叶宝娟等(2013)在变化系数模型下推出合成信度元分析点估计及区间估计的方法。叶宝娟和温忠麟(2013b)还比较了 α 系数的 7 种区间估计方法。

屠金路等(2010)推导出一个多维测验合成分数的信度估计公式。叶宝娟和温忠麟(2012b)探讨了如何用 Delta 法估计多维测验合成信度及其置信区间。杨强等(2014a)比较了 Delta 法与 Bootstrap 法在估计多维测验合成信度的置信区间中的表现。王孟成和叶宝娟(2014)展示了用 *Mplus* 软件实现多维量表各种信度的计算。

在决定将多维测验分数合并成测验总分时, 应当考虑测验同质性高低, 一般认为同质性系数超过 0.5, 合成总分才有意义(温忠麟 等, 2018)。叶宝娟和温忠麟(2012a)探讨了如何用 Delta 法估计多维测验的同质性系数的置信区间, 并与 Bootstrap 法计算的结果进行比较, 发现两者差异很小。顾红磊和温忠麟(2014)发现项目表述效应(由项目表述方式的差异引起的与测量内容无关的系统变异)对自陈测验的同质性系数有影响。

叶宝娟等(2012)介绍了 4 种追踪研究中测验信度的估计方法。叶宝娟和温忠麟(2013c)用两水平验证性因子分析导出一个信度公式进行两水平单维测验的信度估计。田雪垠等(2019)使用 *Mplus* 软件展示了如何利用两水平验证性因子分析计算两水平多维测验的信度估计。

对其他信度的研究包括评分者信度(孙晓敏, 张厚粲, 2005)、质性研究中的编码者信度(徐建平, 张厚粲, 2005)、信度的概括化(关丹丹, 张厚粲, 2004)。此外, 还有其他讨论信度分析的论文(安胜利, 陈平雁, 2001; 顾海根, 李超, 2005; 关丹丹 等, 2005; 关守义, 2009; 焦璨 等, 2008; 刘拓, 戴晓阳, 2011), 有兴趣的读者可以参考。

3.3 中介效应分析

中介效应分析是用来理解自变量 X 对因变量 Y 的影响是如何通过中介变量 M 实现的。温忠麟、张雷等(2004)率先引入简单中介模型并讨论其检验步骤, 推动了国内中介效应分析的方法研究和应用。中介效应的方法学研究主要从 4 个方面展开：检验方法; 中介效应量; 中介模型的拓展; 类别变量的中介效应检验。

中介效应的检验方法研究始于逐步法, 依次检验法和 Sobel 法都可以是逐步法中的关键一步(温忠麟, 张雷 等, 2004)。方杰、张敏强和李晓鹏(2011)综述了后来发展的乘积分布法、Bootstrap 法和贝叶斯法(又称马尔科夫链蒙特卡罗法), 方杰和张敏强(2012, 2013)用模拟方法比较了这三种方法在简单中介分析中的表现, 以及参数和非参数 Bootstrap 法在简单中介分析中的表现。温忠麟和叶宝娟(2014a)提出了一个包含依次检验和 Bootstrap 法的中介效应检验程序。方杰和温忠麟(2018a)比较了 Monte Carlo 法、参数 Bootstrap 法和贝叶斯法在 2-1-1 多层中介分析中的表现。

对于中介效应量, 方杰等(2012)介绍了 4 种中介效应的效应量, 即中介效应占总效应的比例 P_M、中介效应和直接效应的比 R_M、以及两个定义比较复杂的 R^2_{med} 和 κ^2。温忠麟等(2016)明确指出, R^2_{med} 和 κ^2 没有单调性, 不适合作为中介效应量, 他们建议报告 ab 的标准化估计值, 当 ab 和 c 符号一致时才能使用 P_M 作为中介效应量。

简单中介模型被拓展到多重中介模型、多层中介模型、中介与调节的整合模型和纵向数据的

中介模型。柳士顺和凌文辁(2009)介绍了多重中介效应模型, 并用 Sobel 法进行显变量的多重中介模型分析。方杰、温忠麟、张敏强和孙配贞(2014)建议通过增加辅助变量, 利用 Bootstrap 法进行潜变量的多重中介模型分析。方杰等(2010)介绍了显变量的多层中介效应模型分析方法。方杰、温忠麟、张敏强和任皓(2014)进一步介绍了如何利用多层结构方程模型进行潜变量的多层中介效应模型分析。刘红云等(2011)发现当数据符合 1-1-1 随机中介效应模型时, 使用 1-1-1 固定中介效应模型进行多层中介分析将导致错误的中介效应估计。

有调节的中介是指中介过程(X→M→Y)受到调节变量 Z 的影响, 即中介效应大小会随着调节变量 Z 的不同取值而变化。温忠麟等(2006)讨论了如何用依次检验法检验后半路径受到调节的中介模型。温忠麟和叶宝娟(2014b)评述了 3 种有调节的中介模型的分析方法: 依次检验法、系数乘积的区间检验法和中介效应差异检验法。方杰、张敏强等(2014)介绍了如何用系数乘积的区间检验法(贝叶斯法)进行有调节的中介模型检验。方杰和温忠麟(2018b)进一步阐述了如何使用潜调节结构方程方法进行潜变量的有调节的中介模型检验。

有中介的调节模型意味着自变量 X 对因变量 Y 的效应受到调节变量 Z 的影响, 而调节效应(至少部分地)通过中介变量 M 对因变量 Y 起作用。温忠麟等(2006)探讨了如何用依次检验法检验有中介的调节模型。叶宝娟和温忠麟(2013a)系统地评述了检验有中介的调节模型的 5 种方法。另外, 刘国芳等(2018)介绍了基于交叉滞后面板模型、潜变量增长模型和潜变量分数变化模型的纵向中介效应检验。

如果自变量为 k 个类别($k \geqslant 3$), 中介变量和因变量为连续变量的中介效应检验, 方杰等(2017)建议将多类别自变量进行虚拟编码, 考察 $k-1$ 个相对中介效应是否显著。王阳和温忠麟(2018)介绍了两水平被试内设计(即自变量是二分类别变量(如实验组和控制组), 中介变量和因变量是各重复测量二次的连续变量)的中介效应检验。如果中介变量和(或)因变量为类别变量, 应该用 Logistic 回归取代通常的线性回归进行中介效应检验(刘红云 等, 2013; 方杰 等, 2017)。刘红云等(2013)考察了因变量类别数对中介效应检验结果的影响, 发现随着类别数的增加, Logistic 回归与通常线性回归的差别越来越小。方杰等(2017)介绍了中介变量或(和)因变量为类别变量的中介效应检验的发展过程(即尺度统一的过程)。此外, 还有其他讨论中介效应的方法论文(卢谢峰, 韩立敏, 2007; 杜岸政 等, 2014; 甘怡群, 2014; 江程铭, 李纾, 2015), 有兴趣的读者可以参考。

3.4　效应量和检验力

效应量(又称效果量)和检验力(又称统计功效)都和假设检验问题密切相关, 因此将假设检验问题在此一并讨论。吴艳和温忠麟(2011b)讨论了检验力与两类错误率和效应量的关系, 提供了一个可操作的统计分析流程。吕小康(2012)介绍了假设检验思想的提出者 Fisher 与 Neyman-Pearson 在假设检验问题上的分歧。仲晓波的系列研究述评了假设检验受到的批评和争议, 并建议用效应值置信区间表示心理学实验数据分析的结果, 这样能弥补传统假设检验的备择假设不可证伪以及只能提供定性结论的不足(仲晓波 等, 2008; 仲晓波, 2010a; 2010b; 2015; 2016)。胡传鹏等(2016)指出, 过度依赖虚无假设检验是心理学研究可重复性不高的深层原因之一(但也有非统计学根源, 骆大森, 2017)。刘佳等(2018)建议完善虚无假设检验流程来提高心理学研究的可重复性。胡传鹏等(2018)建议将贝叶斯因子(对零假设 H_0 或者备择假设 H_1 的支持程度)作为一种零假设检验的替代和(或)补充方法。

权朝鲁(2003)介绍了效应量的意义和常见效应量的测定方法。温忠麟等(2016)讨论了效应量应具有的性质, 即与测量单位无关、单调性和不受样本容量影响, 还论述了效应量大小和显著性的关系。效应量的定义众多, 郑昊敏等(2011)讨论了三类效应量的选用, 包括差异类、相关类和组重叠类。卢谢峰等(2011)探讨了常用的差异型效应量指标和关联强度型效应量指标的计算公式及使用条件, 并说明了关联强度型效应量指标在 SPSS 软件中的操作。此外, 蒲显伟(2016)也介绍和讨论了效应量的定义、分类、计算方法和解释标准。

在实验研究中效应量常与检验力一起讨论。胡竹菁和戴海琦的系列论文讨论了心理学实验研究中常用检验(Z 检验、t 检验、方差分析和 χ^2 检验)的检验力和效应量的计量原理和方法(胡竹菁, 2010; 胡竹菁, 戴海琦, 2011, 2017)。续志琦和辛自强(2018)介绍了单被试实验的非重叠法效应量

指标。刘铁川等(2019)介绍了方差分析效应量新指标 η_c^2。王珺等(2019)以 Cohen's d 和 η^2 为例，介绍了效应量置信区间的原理和软件实现。赵礼和王晖(2019)介绍了先验检验力分析和后验检验力分析，并展示了如何进行多层数据(即多水平数据)的检验力分析。

3.5 纵向(追踪)数据分析

纵向数据分析的研究有 3 个方面：纵向数据分析模型的发展；纵向数据收集方法的发展；纵向数据的缺失值处理。

刘红云和孟庆茂(2003)率先介绍了如何利用多层线性模型和潜变量增长模型进行纵向数据分析。盖笑松和张向葵(2005)介绍了多层线性模型在纵向研究中的作用。胥彦和李超平(2019)讨论了追踪研究设计的特征、类型、实施步骤、关键点和常用的统计分析模型(包括潜变量增长模型、潜变计分模型、交叉滞后模型)。刘红云(2007)介绍了能描述不同群体不同发展趋势特征的潜变量混合增长模型。潜变量混合增长模型是潜变量增长模型和潜在类别模型的整合，使得纵向研究由注重总体平均趋势的发展过渡到综合考虑总体平均趋势和个体发展的差异上。刘源等(2013)讨论了能描述发展趋势不连续(即存在多个发展阶段)的多阶段增长模型，多阶段增长模型可以解决发展趋势中具有转折点的情形。王婧等(2017)介绍了可同时处理发展过程中的阶段性及群体异质性的多阶段混合增长模型。多阶段混合增长模型是潜变量混合增长模型和多阶段增长模型的整合。刘源等(2014)考察了潜在类别距离、样本量、发展形态等因素对多阶段混合增长模型的模型选择和参数估计的影响。王碧瑶等(2015)建议当研究对象的外显变量和潜在变量都是类别数据时，使用潜在转变模型进行纵向数据分析。潜在转变模型根据作答模式将个体分为不同的潜在类别，并允许个体所属类别随时间的发展发生变化。

数据收集方法的发展包括密集追踪法和加速追踪设计。密集追踪方法是对经验取样法、生态瞬时评估、即时数据获取、日记法等方法的统称(段锦云，陈文平，2012；唐文清 等，2020)。密集追踪法是按一定的程序在现实情境下对被试进行密集测量，获得几十甚至上百个测量点的追踪数据收集方法(唐文清 等，2020)。邵华和吕晓峰(2011)介绍了生态瞬时评估的特点、发展过程和研究设计。段锦云和陈文平(2012)介绍了经验取样法的概念、实施过程、数据处理及其应用。焦璨、熊敏平和张敏强(2010)介绍了函数型数据分析的原理及其应用，函数型数据分析适用于密集追踪数据的分析。唐文清等(2020)考察了利用时变效应模型进行密集追踪数据分析时，样本量和惩罚样条法的节点数对分析结果的影响。唐文清和张敏强等(2014)介绍了加速追踪设计方法，即选择相邻多个群组同时进行短期的追踪研究，获得在测量上有重叠的多个群组追踪数据。

纵向数据的收集过程经常出现数据流失的现象。叶素静等(2014)建议使用选择模型法和模式混合模型法进行缺失值处理。陈楠和刘红云(2015)、张杉杉等(2017)比较了在潜变量增长模型下，不同缺失值处理方法对模型增长参数估计精度及其标准误估计的影响。

此外，郭璐等(2005)介绍了如何用函数回归方法进行纵向数据处理，唐文清、方杰等(2014)对 1982~2012 年发表在 10 本心理学刊物的追踪研究方法的应用论文进行了文献分析。

3.6 调节效应分析

调节效应分析是用来理解自变量 X 对因变量 Y 的影响是如何随着调节变量 Z 而变化的。温忠麟等(2005)首次阐明显变量的调节分析方法并与中介分析进行了比较。方杰等(2015)澄清了均值中心化在调节效应检验中的作用，并讨论了简单斜率检验的两种方法，即选点法和 Johnson-Neyman 法。

潜变量调节效应建模是调节效应研究的核心议题，包括 3 个方面的研究：从有均值的潜调节模型发展到无均值结构的潜调节模型；从需要乘积指标的潜调节模型发展到无需乘积指标的潜调节模型；潜变量调节效应模型的标准化估计。

温忠麟等(2003)评述了潜变量调节效应分析的 4 种方法，包括化潜为显做回归分析、两步最小二乘回归、分组结构方程分析、加入乘积项的结构方程建模。其中，加入乘积项的结构方程模型需要均值结构，影响了该方法的推广应用，吴艳等(2009)提出了一个无需均值结构的潜变量交互效应建模方法，极大地简化了模型。温忠麟和吴艳(2010)系统回顾了乘积指标的潜调节建模方法的简化进程，建议使用无约束方法建立无均值结构的模型。温忠麟等(2013)介绍了无需产生乘积指标的潜变量调节效应分析方法，包括潜调节结

构方程法(LMS)和准极大似然估计法(QML)。方杰等(2018)利用潜调节结构方程法阐述了如何进行多层调节效应分析。

对于潜变量调节效应模型的标准化估计研究，温忠麟等(2008)推导出一个标准化公式来计算潜变量调节效应模型的标准化解，其公式已被 *Mplus* 8.2 以上版本采用。吴艳等(2011)推导出无均值结构的标准化模型，其标准化解与前述的一致。吴艳等(2014)还讨论了潜变量调节效应标准化估计的检验问题。

3.7 探索性因子分析

探索性因子分析的研究主要集中在因子抽取和因子旋转方法。孔明等(2007)介绍了如何用平行分析法来确定抽取的因子数目。高石等(2015)介绍了主轴方法下的平行分析。沐守宽和顾海根(2011)概述了因子抽取的 6 种方法(包括传统的碎石检验法和上文提到的平行分析法)。上述方法都属于因子确证的数理取向，范会勇(2015)还介绍了因子确证的哲理取向和心理趋向。李育辉等(2009)介绍了因子旋转的目标旋转法，并用它评估探索性因子分析结果在因子结构水平上是否具有跨群体的一致性问题。焦璨、吴利和张敏强(2010)讨论了利用目标旋转进行 Bootstrap 探索性因子分析。

此外，刘红云和孟庆茂(2002b)考察因子间相关和因子载荷对探索性因子分析结果的影响。范津砚等(2003)对 1991~2000 年发表在《心理学报》和《心理科学》上的使用探索性因子分析的文章进行了文献分析。还有其他讨论探索性因子分析的论文(胡中锋，莫雷，2002；秦浩，陈景武，2006；孙晓军，周宗奎，2005)，有兴趣的读者可以参考。

3.8 潜在类别模型

潜在类别模型的研究主要从以下 3 个方面开展：探测潜在类别结构的方法；分类的精确性；回归混合模型，即纳入了协变量的潜在类别模型。

探测潜在类别结构的方法有 3 种，即潜在类别分析、潜在剖面分析和类别测量学分析。张洁婷等(2010)率先介绍了潜在类别分析及其在心理学中的应用。张洁婷等(2012)介绍了潜变量的数据类型和空间特征的研究现状，建议通过模型(如潜在类别模型、混合因子模型和因子模型)的嵌套比较来确定潜变量的类型。吕浥尘和赵然(2018)介绍了群组发展模型(一种潜在类别分析方法)在干预研究中的应用。马文超等(2014)比较了 K-means 法、潜在类别模型和混合 Rasch 模型在二分外显变量情境下的聚类效果。外显变量是连续变量的潜在类别分析则称为潜在剖面分析。尹奎等(2020)介绍了潜在剖面分析及其在组织行为领域的应用。任芬等(2013)介绍了在精神病学中常用的类别测量学分析法，该分析法通过探测一组外显变量背后的潜变量是连续的还是类别的，来确定群体中是否存在异质群体。

王孟成、邓俏文和毕向阳等(2017)考查了潜在剖面分析中，样本量、潜在类别数目、类别距离和指标个数对分类精确性指数熵的影响。王孟成和毕向阳(2018)介绍了回归混合模型的各种分析方法。张洁婷等(2017, 2019)比较了纳入式分类分析法、三步法等在回归混合模型分析中的表现，并给出了分析流程。

3.9 共同方法偏差

温忠麟等(2018)将共同方法偏差列为问卷数据建模之前的必备工作之一，许多期刊也都要求检验共同方法偏差是否严重。共同方法偏差的研究主要从以下 3 个方面开展：共同方法偏差和共同方法变异的关系；共同方法偏差的影响；共同方法偏差的检验和控制方法。

熊红星等(2012)认为共同方法偏差是共同方法变异导致的偏差性结果，因此使用“共同方法变异”更好。朱海腾和李川云(2019)也对共同方法变异与共同方法偏差进行了辨析。当存在共同方法变异时，出现共同方法偏差的概率增大，但尚未发现二者有稳定的强对应关系。

杜建政等(2005)介绍了共同方法偏差对测评的影响，一是表现在行为测评的数据上，二是表现在测评数据之间的相关上。熊红星等(2013)进一步指出，方法变异以两种方式影响观测数据，即改变所研究的潜在结构和歪曲测量的过程。朱海腾和李川云(2019)指出共同方法变异普遍存在，数据来源、测量时间、问卷设计等因素可导致共同方法偏差，已有研究者从测量误差和非共同方法变异的制约作用等角度说明无需过度担忧共同方法变异的消极影响。

共同方法偏差的控制可分为程序控制和统计控制。周浩和龙立荣(2004)率先介绍了共同方差偏差的检验和控制方法，推动了国内问卷类研究对共同方法偏差进行检验与控制。他们着重介绍了因子分析法、偏相关法、潜在的误差变量控制法、

结构方程模型法等。熊红星等(2012)介绍了统计控制的各种模型, 认为控制潜在方法因子途径(包括双因子模型和多质多法)是统计控制法中最重要的一类方法。温忠麟等(2018)建议使用潜在的误差变量控制法, 即建立双因子模型, 在原有设计的因子基础上加上一个方法因子作为全局因子, 并比较加入方法因子前后模型的拟合情况。张缨斌和王烨晖(2019)介绍了反应风格(共同方法偏差的来源之一)的定义、类型、危害、测量和统计控制方法。汤丹丹和温忠麟(2020)指出一个好的共同方法偏差检验法应该具有三个特征, 即符合共同方法变异的数学模型、评价标准不受非共同方法变异来源的影响和对共同方法变异、共同方法偏差的变化敏感。利用这三个特征系统评述了检验共同方法偏差的 Harman 单因子法、双因子法和验证性因子分析标签法。

3.10 多层线性模型

刘红云和孟庆茂(2002a)从多层线性模型的基本假设入手, 系统地介绍了模型的参数估计和假设检验的方法, 推动了国内多层线性模型研究的发展。多层线性模型的发展主要体现在 3 个方面: 个体层次的评价聚合到群体层次时的聚合适当性检验; 因变量是类别的多层线性模型; 潜变量的多层线性模型, 即多层结构方程模型。

多层数据分析中, 常需要将个体层次的评分聚合到群体层次, 为确保聚合后的分数具有充分代表性, 需要进行聚合适当性检验。徐晓锋和刘勇(2007)介绍了评分者一致性指标 γ_{WG}, 以及评分者一致性与评分者信度的区别。杨建锋和王重鸣(2008)介绍了 10 种类内相关系数(Intraclass Correlation Coefficients, ICCs), 其中 ICC (1)和 ICC (2)用于判断数据是否适合进行多层分析。朱海腾(2020)建议将 γ_{WG} 作为聚合适当性指标, ICC (1)和 ICC (2)分别作为效度、信度指标。

李雪燕和辛涛(2006)以学生嵌套于班级为例, 介绍了因变量是二分类别变量的多层线性模型分析, 即多层二项 logit 模型。谢美华等(2013)介绍了因变量是多类别变量的多层线性模型分析, 即多层多项 logit 模型。

多层结构方程模型能同时解决数据嵌套和考虑测量误差的问题。方杰、邱皓政和张敏强(2011)介绍了如何用多层结构方程模型进行情境效应分析。多层结构方程模型和多层线性模型用于中介分析、调节分析、纵向数据分析等内容详见本文相应部分阐述。此外, 方杰等(2013)还对 2002~2011 年发表在 10 本心理学刊物上的多层线性模型的应用论文进行了文献分析。

4 结语

进入 21 世纪的 20 年来, 国内心理统计研究取得了长足的进步, 研究领域不断拓展(已形成 10 个研究热点), 研究的深度不断增加(有 4 类研究的文篇数都超过 20 篇), 有三个心理统计方法研究团队的论文较多, 这三个团队及其毕业生的工作, 撑起了国内心理统计方法研究的半壁江山。这期间发表的心理统计方法类论文(三个团队带头人中有合作的论文, 计入排名靠前的作者团队), 温忠麟团队共有 48 篇, 研究主要涉及中介效应、调节效应、结构方程模型、信度和效应量; 张敏强团队共有 27 篇, 研究主要涉及纵向研究、潜在类别分析、中介效应; 刘红云团队共有 17 篇, 研究主要涉及结构方程模型、纵向研究和中介效应。在心理统计的研究中, 各个研究热点相互融合中共同发展, 例如中介与调节的整合、中介或者调节与结构方程模型、多层线性模型的整合。还有一些新兴的研究领域也受到关注, 例如社会网络分析(蔡玉清 等, 2020; 黄菲菲, 张敏强, 2016; 焦璨 等, 2014; 马绍奇 等, 2011; 徐伟 等, 2011)。

需要说明的是, 本文仅对 10 个研究热点的国内心理统计研究进行了文献分析, 至于各个研究热点问题更详细的内容分析需专文讨论。除了这 10 个研究热点和社会网络分析外, 国内心理统计研究还涉及缺失值处理(沐守宽, 周伟, 2011; 王孟成, 邓俏文, 2016)、不同形式的回归模型(唐杰 等, 2011; 徐芃 等, 2015; 张沥今 等, 2020; 张珊珊 等, 2012)、社会关系模型(郭素然, 伍新春, 2013; 徐桃 等, 2015; 张宏宇 等, 2007)、贝叶斯统计推断(钟建军 等, 2017)、趋中回归与超常分数重现概率(温忠麟, 侯杰泰, 2003)、实验数据的随机化检验(吕小康, 付英涛, 2019)、均匀设计(范会勇 等, 2009)、整合数据分析(唐文清 等, 2012)、粗糙集和神经网络分析(余嘉元, 2005; 2007; 2008)、多维尺度分析(赵广平, 2015)、优势分析(谢宝国, 龙立荣, 2006)、关联规则方法(王冬燕 等, 2013)、计划缺失设计(王孟成, 叶浩生, 2014)、因果推论与分析(温忠麟, 2017)、Schuirmann 无差异

表 2　2001~2020 年发表的被引超过 300 次的心理统计方法论文(按被引量排序)

篇名	期刊	发表年份	第一作者	被引次数
中介效应检验程序及其应用	心理学报	2004	温忠麟	9284
调节效应与中介效应的比较和应用	心理学报	2005	温忠麟	6242
中介效应分析: 方法和模型发展	心理科学进展	2014	温忠麟	4651
共同方法偏差的统计检验与控制方法	心理科学进展	2004	周　浩	3054
结构方程模型检验: 拟合指数与卡方准则	心理学报	2004	温忠麟	2467
有中介的调节变量和有调节的中介变量	心理学报	2006	温忠麟	1294
有调节的中介模型检验方法: 竞争还是替补?	心理学报	2014	温忠麟	992
中介效应的检验方法和效果量测量: 回顾与展望	心理发展与教育	2012	方　杰	587
结构方程建模中的题目打包策略	心理科学进展	2011	吴　艳	566
共同方法变异的影响及其统计控制途径的模型分析	心理科学进展	2012	熊红星	397
基于结构方程模型的多重中介效应分析	心理科学	2014	方　杰	395
有中介的调节模型检验方法: 甄别和整合	心理学报	2013	叶宝娟	347
多重中介模型及其应用	心理科学	2009	柳士顺	310
基于多元回归的调节效应分析	心理科学	2015	方　杰	304

注：截至 2021 年 2 月 10 日，中国知网查询结果。

检验(王力 等, 2005)、多元总体均值差异显著性检验(田晓明, 傅珏生, 2005)、差数显著性 t 检验与元分析方法的模拟对比(郭春彦 等, 2002)、定量研究和应用统计方法应注意的问题(陈启山, 2006; 单志艳, 孟庆茂, 2002)等议题。

当然，国内心理统计的研究仍存在不足。首先，综述研究占了较大的比例(如《心理科学进展》有 55 篇)，将国外最新的统计进展引入国内是及时的，但采用数据模拟进行的心理统计研究只有 40 篇(心理学报 22 篇，心理科学 11 篇，心理学探新 7 篇，其中 32 篇都是近 10 年发表的)，原创性的心理统计研究还有较大的进步空间。其次，国内心理统计的科研力量相比心理学的其他分支还相对薄弱，虽然涌现出方杰(18 篇，主要研究中介效应和调节效应)、叶宝娟(17 篇，主要研究测验信度)等新秀(未列出在国外期刊发文的中国籍优秀研究者)，但与心理学对统计的需求相比还是不够的，研究力量有待加强。相信随着国家对中文期刊论文的重视，会吸引更多的高质量心理统计方法论文发表在国内期刊上。

国内 2001~2020 年发表的心理统计学论文对心理学、行为学、管理学等社科领域的实证研究有非常大的影响，截至 2021 年 2 月 10 日，被引超过 300 次的就有 14 篇(表 2)。随着心理统计方法的发展，研究者可以更好地分析各种心理数据，更好地理解和解释各种复杂的心理现象。

参考文献

安胜利, 陈平雁. (2001). 量表的信度及其影响因素. *中国临床心理学杂志*, *9*(4), 315–318.

白新文, 陈毅文. (2004). 测量等价性的概念及其判定条件. *心理科学进展*, *12*(2), 231–239.

卞冉, 车宏生, 阳辉. (2007). 项目组合在结构方程模型中的应用. *心理科学进展*, *15*(3), 567–576.

蔡玉清, 董书阳, 袁帅, 胡传鹏. (2020). 变量间的网络分析模型及其应用. *心理科学进展*, *28*(1), 178–190.

陈楠, 刘红云. (2015). 基于增长模型的非随机缺失数据处理: 选择模型和极大似然方法. *心理科学*, *38*(2), 446–451.

陈启山. (2006). 心理学研究中应用统计方法应注意的几个问题. *心理与行为研究*, *4*(3), 200–206.

陈宇帅, 温忠麟, 顾红磊. (2015). 因子混合模型: 潜在类别分析与因子分析的整合. *心理科学进展*, *23*(3), 529–538.

崔丽霞, 郑日昌. (2001). 20 年来我国心理学研究方法的回顾与反思. *心理学报*, *33*(6), 85–91.

杜岸政, 古纯文, 丁桂凤. (2014). 心理学研究中的中介效应分析意义及方法评述. *中国心理卫生杂志*, *28*(8), 578–583.

杜建政, 赵国祥, 刘金平. (2005). 测评中的共同方法偏差. *心理科学*, *28*(2), 420–422.

段锦云, 陈文平. (2012). 基于动态评估的取样法: 经验取

样法. *心理科学进展*, *20*(7), 1110–1120.
范会勇. (2015). 试论因子确证的三种取向及其整合. *心理学探新*, *35*(3), 257–260.
范会勇, 杨新国, 张进辅. (2009). 均匀设计在心理实验中的应用探讨. *心理科学进展*, *17*(1), 233–239.
范津砚, 叶斌, 章震宇, 刘宝霞. (2003). 探索性因素分析——最近 10 年的评述. *心理科学进展*, *11*(5), 579–585.
方杰, 邱皓政, 张敏强. (2011). 基于多层结构方程模型的情境效应分析——兼与多层线性模型比较. *心理科学进展*, *19*(2), 284–292.
方杰, 邱皓政, 张敏强, 方路. (2013). 我国近十年来心理学研究中 HLM 方法的应用述评. *心理科学*, *36*(5), 1194–1200.
方杰, 温忠麟. (2018a). 三类多层中介效应分析方法比较. *心理科学*, *41*(4), 962–967.
方杰, 温忠麟. (2018b). 基于结构方程模型的有调节的中介效应分析. *心理科学*, *41*(2), 453–458.
方杰, 温忠麟, 梁东梅, 李霓霓. (2015). 基于多元回归的调节效应分析. *心理科学*, *38*(3), 715–720.
方杰, 温忠麟, 吴艳. (2018). 基于结构方程模型的多层调节效应. *心理科学进展*, *26*(5), 781–788.
方杰, 温忠麟, 张敏强. (2017). 类别变量的中介效应分析. *心理科学*, *40*(2), 471–477.
方杰, 温忠麟, 张敏强, 任皓. (2014). 基于结构方程模型的多层中介效应分析. *心理科学进展*, *22*(3), 530–539.
方杰, 温忠麟, 张敏强, 孙配贞. (2014). 基于结构方程模型的多重中介效应分析. *心理科学*, *37*(3), 735–741.
方杰, 张敏强. (2012). 中介效应的点估计和区间估计：乘积分布法、非参数 Bootstrap 和 MCMC 法. *心理学报*, *44*(10), 1408–1420.
方杰, 张敏强. (2013). 参数和非参数 Bootstrap 方法的简单中介效应分析比较. *心理科学*, *36*(3), 722–727.
方杰, 张敏强, 顾红磊, 梁东梅. (2014). 基于不对称区间估计的有调节的中介模型检验. *心理科学进展*, *22*(10), 1660–1668.
方杰, 张敏强, 李晓鹏. (2011). 中介效应的三类区间估计方法. *心理科学进展*, *19*(5), 765–774.
方杰, 张敏强, 邱皓政. (2010). 基于阶层线性理论的多层级中介效应. *心理科学进展*, *18*(8), 1329–1338.
方杰, 张敏强, 邱皓政. (2012). 中介效应的检验方法和效果量测量: 回顾与展望. *心理发展与教育*, *28*(1), 105–111.
方平, 熊端琴, 蔡红. (2001). 结构方程在心理学研究中的应用. *心理科学*, *24*(4), 406–408.
方平, 熊端琴, 曹雪梅. (2002). 结构方程模式的发展与应用. *心理科学进展*, *10*(3), 270–279.
盖笑松, 张向葵. (2005). 多层线性模型在纵向研究中的运用. *心理科学*, *28*(2), 429–431.
甘怡群. (2014). 中介效应研究的新趋势——研究设计和数据统计方法. *中国心理卫生杂志*, *28*(8), 584–585.
高石, 杨莉萍, 白福宝. (2015). 探索性因子分析中主轴法下的平行分析. *心理学探新*, *35*(5), 471–473.
顾海根, 李超. (2005). 同质信度多种指标的比较研究. *心理科学*, *28*(5), 1196–1198.
顾红磊, 温忠麟. (2014). 项目表述效应对自陈量表信效度的影响——以核心自我评价量表为例. *心理科学*, *37*(5), 1245–1252.
顾红磊, 温忠麟. (2017). 多维测验分数的报告与解释: 基于双因子模型的视角. *心理发展与教育*, *33*(4), 504–512.
顾红磊, 温忠麟, 方杰. (2014). 双因子模型: 多维构念测量的新视角. *心理科学*, *37*(4), 973–979.
关丹丹, 张厚粲. (2004). 信度的再认识与信度概括化研究. *心理科学*, *27*(2), 445–448.
关丹丹, 张厚粲, 李中权. (2005). 差异分数的信度分析. *心理科学*, *28*(1), 161–163.
关守义. (2009). 克龙巴赫 α 系数研究述评. *心理科学*, *32*(3), 685–687.
郭春彦, 朱滢, 王全珍. (2002). 差数显著性 *t* 检验与元分析方法的模拟对比. *心理学报*, *34*(2), 46−50.
郭璐, 沈模卫, 关忻. (2005). 函数回归方法在肢体运动轨迹预测中的应用. *心理科学*, *28*(3), 660−662.
郭庆科, 李芳, 陈雪霞, 王炜丽, 孟庆茂. (2008). 不同条件下拟合指数的表现及临界值的选择. *心理学报*, *40*(1), 109–118.
郭庆科, 王炜丽, 陈雪霞, 韩丹. (2007). 验证性因素分析中模型拟合的判断. *心理学探新*, *27*(4), 83–87.
郭素然, 伍新春. (2013). 社会关系模型在家庭研究中的应用. *心理科学进展*, *21*(3), 561–570.
胡传鹏, 孔祥祯, Eric-Jan Wagenmakers, Alexander Ly, 彭凯平. (2018). 贝叶斯因子及其在 JASP 中的实现. *心理科学进展*, *26*(6), 951–965.
胡传鹏, 王非, 过继成思, 宋梦迪, 隋洁, 彭凯平. (2016). 心理学研究中的可重复性问题: 从危机到契机. *心理科学进展*, *24*(9), 1504–1518.
胡中锋, 莫雷. (2002). 论因素分析方法的整合. *心理科学*, *25*(4), 474–475.
胡竹菁. (2010). 平均数差异显著性检验统计检验力和效果大小的估计原理与方法. *心理学探新*, *30*(1), 68–73.
胡竹菁, 戴海琦. (2011). 方差分析的统计检验力和效果大小的常用方法比较. *心理学探新*, *31*(3), 254–259.
胡竹菁, 戴海琦. (2017). 心理学实验研究的效果大小. *心理学探新*, *37*(1), 70–77.
黄菲菲, 张敏强. (2016). 社会网络分析中缺失数据的处理方法. *心理技术与应用*, *4*(8), 456–464.
江程铭, 李纾. (2015). 中介分析和自举(Bootstrap) 程序应用. *心理学探新*, *35*(5), 458–463.
焦璨, 黄泽娟, 张敏强, 吴利, 王宣承. (2010). 心理研究中统计方法应用的元分析——以《心理学报》《心理科学》(1998-2008)统计方法应用为例. *心理科学*, *33*(1), 48–54.
焦璨, 吴换杰, 黄玥娜, 黄菲菲, 张敏强. (2014). 网络自相关模型在心理学研究中的应用 ——以同群效应、学习动机对青少年学业表现的影响为例. *心理学报*,

46(12), 1933–1945.

焦璨, 吴利, 张敏强. (2010). Bootstrap 探索性因素分析方法及其应用. *心理学探新*, *30*(3), 84–90.

焦璨, 熊敏平, 张敏强. (2010). 心理学研究数据类型与统计方法——谈函数型数据分析的引入. *心理科学进展*, *18*(8), 1314–1320.

焦璨, 张敏强, 黄庆均, 张文怡, 黎光明. (2008). 非正态分布测量数据对克隆巴赫信度 α 系数的影响. *应用心理学*, *14*(3), 276–281.

李冲, 刘红云. (2011). 等级数据的测量不变性检验及影响因素模拟研究. *心理科学*, *34*(6), 1482–1487.

李旭, Gail Huon, 钱铭怡. (2001). 在临床心理学领域运用结构方程模型的思路与步骤. 中国临床心理学杂志, *9*(2), 149–152.

李雪燕, 辛涛. (2006). 二分数据的多层线性模型: 原理与应用. *心理发展与教育*, *22*(4), 97–102.

李育辉, 黄飞. (2010). 成对数据分析之行动者-对象互依性模型(APIM). *心理科学进展*, *18*(8), 1321–1328.

李育辉, 黄飞, 张建新. (2009). 探索性因素分析、目标旋转与因素结构的一致性. *心理科学进展*, *17*(5), 1094–1101.

梁莘娅, 杨艳云. (2016). 当结构假设和分布假设不满足时的验证性因子分析: 稳健极大似然估计和贝叶斯估计的比较研究(英文). *心理科学*, *39*(5), 1256–1267.

刘畅, 伍新春. (2017). 主客体互倚性的成对模式及其检验. *心理发展与教育*, *33*(1), 105–112.

刘国芳, 程亚华, 辛自强. (2018). 作为因果关系的中介效应及其检验. *心理技术与应用*, *6*(11), 665–676.

柳恒超, 许燕, 王力. (2007). 结构方程模型应用中模型选择的原理和方法. *心理学探新*, *27*(1), 75–78.

刘红云. (2007). 如何描述发展趋势的差异: 潜变量混合增长模型. *心理科学进展*, *15*(3), 539–544.

刘红云. (2008). α 系数与测验的同质性. *心理科学*, *31*(1), 185–188.

刘红云, 李冲, 张平平, 骆方. (2012). 分类数据测量等价性检验方法及其比较:项目阈值(难度)参数的组间差异性检验. *心理学报*, *44*(8), 1124–1136.

刘红云, 李美娟, 骆方, 李小山. (2012). 单维项目因素分析: CCFA 与 IRT 估计方法的比较. *心理科学*, *35*(2), 441–445.

刘红云, 骆方, 王玥, 张玉. (2012). 多维测验项目参数的估计:基于 SEM 与 MIRT 方法的比较. *心理学报*, *44*(1), 121–132.

刘红云, 骆方, 张玉, 张丹慧. (2013). 因变量为等级变量的中介效应分析. *心理学报*, *45*(12), 1431–1442.

刘红云, 孟庆茂. (2002a). 教育和心理研究中的多层线性模型. *心理科学进展*, *10*(2), 213–219.

刘红云, 孟庆茂. (2002b). 探索性因素分析在测验编制中局限性的模拟实验. *心理科学*, *25*(2), 177–179.

刘红云, 孟庆茂. (2003). 纵向数据分析方法. *心理科学进展*, *11*(5), 586–592.

刘红云, 张月, 骆方, 李美娟, 李小山. (2011). 多水平随机中介效应估计及其比较. *心理学报*, *43*(6), 696–709.

刘佳, 霍涌泉, 陈文博, 解诗薇, 王静. (2018). 心理学研究的可重复性"危机": 一些积极应对策略. *心理学探新*, *38*(1), 86–90.

刘军, 吴维库. (2005). 心理测量平衡性研究与实例. *心理科学*, *28*(1), 170–174.

柳士顺, 凌文辁. (2009). 多重中介模型及其应用. *心理科学*, *32*(2), 433–435.

刘铁川, 王闪闪, 桂雅立. (2019). 方差分析效果大小报告的新指标. *心理学探新*, *39*(3), 238–243.

刘拓, 戴晓阳. (2011). 不拟合被试对测验信、效度的影响. *中国临床心理学杂志*, *19*(6), 743–745.

刘源, 骆方, 刘红云. (2014). 多阶段混合增长模型的影响因素: 距离与形态. *心理学报*, *46*(9), 1400–1412.

刘源, 赵骞, 刘红云. (2013). 多阶段增长模型的方法比较. *心理学探新*, *33*(5), 415–422.

刘玥, 刘红云. (2017). 基于双因子模型的测验总分和维度分的合成方法. *心理学报*, *49*(9), 1234–1246.

卢谢峰, 韩立敏. (2007). 中介变量、调节变量与协变量——概念、统计检验及其比较. *心理科学*, *30*(4), 934–936.

卢谢峰, 唐源鸿, 曾凡梅. (2011). 效应量: 估计、报告和解释. *心理学探新*, *31*(3), 260–264.

骆大森. (2017). 心理学可重复性危机两种根源的评估. *心理与行为研究*, *15*(5), 577–586.

吕小康. (2012). Fisher 与 Neyman-Pearson 的分歧与心理统计中的假设检验争议. *心理科学*, *35*(6), 1502–1506.

吕小康, 付英涛. (2019). 实验数据的随机化检验及 R 语言实现. *心理技术与应用*, *7*(5), 266–275.

吕浥尘, 赵然. (2018). 群组发展模型——干预研究的新方法. *心理学探新*, *38*(1), 91–96.

孔明, 卞冉, 张厚粲. (2007). 平行分析在探索性因素分析中的应用. *心理科学*, *30*(4), 924–925.

马绍奇, 焦璨, 张敏强. (2011). 社会网络分析在心理研究中的应用. *心理科学进展*, *19*(5), 755–764.

马文超, 边玉芳, 郭雯婧, 谢敏. (2014). K-means、潜在类别模型和混合 Rasch 模型的比较. *心理学探新*, *34*(5), 431–436.

麦玉娇, 温忠麟. (2013). 探索性结构方程建模(ESEM): EFA 和 CFA 的整合. *心理科学进展*, *21*(5), 934–939.

孟庆茂, 刘红云. (2002). α 系数在使用中存在的问题. *心理学探新*, *22*(3), 42–47.

沐守宽, 顾海根. (2011). 探索性因素分析因子抽取方法的比较. *心理学探新*, *31*(5), 477–480.

沐守宽, 周伟. (2011). 缺失数据处理的期望-极大化算法与马尔可夫蒙特卡洛方法. *心理科学进展*, *19*(7), 1083–1090.

蒲显伟. (2016). 定量数据分析效应值: 意义, 计算与解释. *心理学探新*, *36*(1), 64–69.

秦浩, 陈景武. (2006). 量表编制过程中应用因子分析容易

忽略的几个问题. *中国心理卫生杂志*, *20*(5), 320-321.
权朝鲁. (2003). 效果量的意义及测定方法. *心理学探新*, *23*(2), 39–44.
任芬, 潘林, 王孟成, 姚树桥, 张建新. (2013). 区分类别与维度: Taxometric 分析法简介. *中国临床心理学杂志*, *21*(5), 786–789.
任学柱, 王腾飞, Karl Schweizer. (2017). 预设路径模型及其在认知心理学研究中的应用. *心理科学进展*, *25*(10), 1675–1681.
单志艳, 孟庆茂. (2002). 心理学中定量研究的几个问题. *心理科学*, *25*(4), 466–467.
邵华, 吕晓峰. (2011). EMA——一种生态主义取向的研究模式. *心理科学*, *34*(5), 1236–1241
孙晓军, 周宗奎. (2005). 探索性因子分析及其在应用中存在的主要问题. *心理科学*, *28*(6), 1440–1442.
孙晓敏, 张厚粲. (2005). 表现性评价中评分者信度估计方法的比较研究——从相关法、百分比法到概化理论. *心理科学*, *28*(3), 646–649.
汤丹丹, 温忠麟. (2020). 共同方法偏差检验: 问题与建议. *心理科学*, *43*(1), 215–223.
唐杰, 林志扬, 莫莉. (2011). 多项式回归与一致性研究: 应用及分析. *心理学报*, *43*(12), 1454–1461.
唐文清, 方杰, 蒋香梅, 张敏强. (2014). 追踪研究方法在国内心理研究中的应用述评. *心理发展与教育*, *30*(2), 216–224.
唐文清, 张敏强, 方杰. (2020). 时变效应模型及在密集追踪数据分析中的应用. *心理科学*, *43*(2), 488–497.
唐文清, 张敏强, 黄宪, 张嘉志, 王旭. (2014). 加速追踪设计的方法和应用. *心理科学进展*, *22*(2), 369–380.
唐文清, 张敏强, 王力田. (2012). 整合数据分析方法在心理学研究中的应用. *心理学探新*, *32*(5), 454–460.
田晓明, 傅珏生. (2005). 多元总体均值差异显著性检验的研究. *心理科学*, *28*(1), 164–165.
田雪垠, 郑蝉金, 郭少阳, 贺冠瑞. (2019). 基于多层验证性因素分析的各种信度系数方法. *心理学探新*, *39*(5), 461–467.
屠金路, 金瑜, 王庭照. (2005). bootstrap 法在合成分数信度区间估计中的应用. *心理科学*, *28*(5), 1199–1200.
屠金路, 王庭照, 金瑜. (2010). 结构方程模型下多因子非同质测量合成分数的信度估计. *心理科学*, *33*(3), 666–669.
王碧瑶, 张敏强, 张洁婷, 胡俊. (2015). 基于转变矩阵描述的个体阶段性发展: 潜在转变模型. *心理研究*, *8*(4), 36–43.
王冬燕, 钱锦昕, 徐海宁, 余嘉元. (2013). 关联规则方法在心理测量中的应用. *心理科学*, *36*(2), 475-478.
王婧, 唐文清, 张敏强, 张文怡, 郭凯茵. (2017). 多阶段混合增长模型的方法及研究现状. *心理科学进展*, *25*(10), 1696–1704.
王珺, 宋琼雅, 许岳培, 贾彬彬, 胡传鹏. (2019). 效应量置信区间的原理及其实现. *心理技术与应用*, *7*(5), 284–296.
王力, 唐云, 李中权. (2005). Schuirmann 无差异检验在心理学研究中的意义与应用. *中国心理卫生杂志*, *19*(5), 316–318.
王孟成, 毕向阳. (2018). 回归混合模型: 方法进展与软件实现. *心理科学进展*, *26*(12), 2272–2280.
王孟成, 邓俏文. (2016). 缺失数据的结构方程建模:全息极大似然估计时辅助变量的作用. *心理学报*, *48*(11), 1489–1498.
王孟成, 邓俏文, 毕向阳. (2017). 潜变量建模的贝叶斯方法. *心理科学进展*, *25*(10), 1682–1695.
王孟成, 邓俏文, 毕向阳, 叶浩生, 杨文登. (2017). 分类精确性指数 Entropy 在潜剖面分析中的表现: 一项蒙特卡罗模拟研究. *心理学报*, *49*(11), 1473–1482.
王孟成, 叶宝娟. (2014). 通过 *Mplus* 计算几种常用的测验信度. *心理学探新*, *34*(1), 48–52.
王孟成, 叶浩生. (2014). 计划缺失设计——通过有意缺失让研究更高效. *心理科学进展*, *22*(6), 1025–1035.
王晓丽, 李西营, 邵景进. (2011). 形成性测量模型: 结构方程模型的新视角. *心理科学进展*, *19*(2), 293–300.
王阳, 温忠麟. (2018). 基于两水平被试内设计的中介效应分析方法. *心理科学*, *41*(5), 1233–1239.
王阳, 温忠麟, 付媛姝. (2020). 等效性检验——结构方程模型评价和测量不变性分析的新视角. *心理科学进展*, *28*(11), 1961–1969.
韦嘉, 张春雨, 赵永萍, 张进辅. (2016). 随机截距因子分析模型在控制条目表述效应中的应用. *心理科学*, *39*(4), 1005–1010.
温聪聪, 伍伟平, 林光杰. (2019). 对齐(Alignment)——一种新的多群组分析法. *心理科学进展*, *27*(1), 181–189.
温涵, 梁韵斯. (2015). 结构方程模型常用拟合指数检验的实质. *心理科学*, *38*(4), 987–994.
温忠麟. (2017). 实证研究中的因果推理与分析. *心理科学*, *40*(1), 200–208.
温忠麟, 范息涛, 叶宝娟, 陈宇帅. (2016). 从效应量应有的性质看中介效应量的合理性. *心理学报*, *48*(4), 435–443.
温忠麟, 侯杰泰. (2003). 心理测验中的趋中回归与超常分数重现概率. *心理学报*, *35*(3), 419–425.
温忠麟, 侯杰泰. (2008). 检验的临界值: 真伪差距多大才能辨别?——评《不同条件下拟合指数的表现及临界值的选择》. *心理学报*, *40*(1), 119–124.
温忠麟, 侯杰泰, 马什赫伯特. (2003). 潜变量交互效应分析方法. *心理科学进展*, *11*(5), 593–599.
温忠麟, 侯杰泰, 马什赫伯特. (2004). 结构方程模型检验: 拟合指数与卡方准则. *心理学报*, *36*(2), 186–194.
温忠麟, 侯杰泰, Herbert W. Marsh. (2008). 结构方程模型中调节效应的标准化估计. *心理学报*, *40*(6), 729–736.
温忠麟, 侯杰泰, 张雷. (2005). 调节效应与中介效应的比较和应用. *心理学报*, *37*(2), 268–274.
温忠麟, 黄彬彬, 汤丹丹. (2018). 问卷数据建模前传. *心*

理科学, *41*(1), 204–210.
温忠麟, 汤丹丹, 顾红磊. (2019). 预测视角下双因子模型与高阶因子模型的一般性模拟比较. *心理学报*, *51*(3), 383–391.
温忠麟, 吴艳. (2010). 潜变量交互效应建模方法演变与简化. *心理科学进展*, *18*(8), 1306–1313.
温忠麟, 吴艳, 侯杰泰. (2013). 潜变量交互效应结构方程: 分布分析方法. *心理学探新*, *33*(5), 409−414.
温忠麟, 叶宝娟. (2011). 测验信度估计: 从 α 系数到内部一致性信度. *心理学报*, *43*(7), 821–829.
温忠麟, 叶宝娟. (2014a). 中介效应分析: 方法和模型发展. *心理科学进展*, *22*(5), 731–745.
温忠麟, 叶宝娟. (2014b). 有调节的中介模型检验方法: 竞争还是替补? *心理学报*, *46*(5), 714–726.
温忠麟, 张雷, 侯杰泰. (2006). 有中介的调节变量和有调节的中介变量. *心理学报*, *38*(3), 448–452.
温忠麟, 张雷, 侯杰泰, 刘红云. (2004). 中介效应检验程序及其应用. *心理学报*, *36*(5), 614–620.
吴瑞林, 涂冬波. (2013). 题目因素分析: 基于 SEM 和基于 IRT 的两类方法. *心理与行为研究*, *11*(1), 124–131.
吴艳, 温忠麟. (2011a). 结构方程建模中的题目打包策略. *心理科学进展*, *19*(12), 1859–1867.
吴艳, 温忠麟. (2011b). 与零假设检验有关的统计分析流程. *心理科学*, *34*(1), 230–234.
吴艳, 温忠麟, 侯杰泰, Herbert W. Marsh. (2011). 无均值结构的潜变量交互效应模型的标准化估计. *心理学报*, *43*(10), 1219–1228.
吴艳, 温忠麟, 李碧. (2014). 潜变量交互效应模型标准化估计中的检验问题. *心理学探新*, *34*(3), 260–264.
吴艳, 温忠麟, 林冠群. (2009). 潜变量交互效应建模: 告别均值结构. *心理学报*, *41*(2), 1252–1259.
谢宝国, 龙立荣. (2006). 优势分析方法及其应用. *心理科学*, *29*(4), 922–925.
谢美华, 黄友泉, 董圣鸿. (2013). 多层多项 Logit 模型: 原理与应用. *心理学探新*, *33*(5), 430–437.
熊红星, 张璟, 郑雪. (2013). 方法影响结果?方法变异的本质、影响及控制. *心理学探新*, *33*(3), 195–199.
熊红星, 张璟, 叶宝娟, 郑雪, 孙配贞. (2012). 共同方法变异的影响及其统计控制途径的模型分析. *心理科学进展*, *20*(5), 757–769.
徐芃, 祁禄, 熊健, 叶浩生. (2015). 定序变量回归模型在心理数据分析中的应用. *心理学报*, *47*(12), 1520–1528.
徐建平, 张厚粲. (2005). 质性研究中编码者信度的多种方法考察. *心理科学*, *28*(6), 1430–1432.
徐霜雪, 俞宗火, 李月梅. (2017). 预测视角下双因子模型与高阶模型的模拟比较. *心理学报*, *49*(8), 1125–1136.
徐桃, 张敏强, 王小婷, 黄兆锋, 焦璨. (2015). 社会关系模型在心理研究中的应用. *心理科学进展*, *23*(3), 520–528.
徐伟, 陈光辉, 曾玉, 张文新. (2011). 关系研究的新取向: 社会网络分析. *心理科学*, *34*(2), 499–504.
徐晓锋, 刘勇. (2007). 评分者内部一致性的研究和应用. *心理科学*, *30*(5), 1175–1178.
胥彦, 李超平. (2019). 追踪研究在组织行为学中的应用. *心理科学进展*, *27*(4), 600–610.
续志琦, 辛自强. (2018). 单被试实验的统计分析: 非重叠法效果量估计. *心理技术与应用*, *6*(2), 89–99.
杨建锋, 王重鸣. (2008). 类内相关系数的原理及其应用. *心理科学*, *31*(2), 434–437.
杨强, 叶宝娟, 温忠麟. (2014a). 两种估计多维测验合成信度置信区间方法比较. *心理学探新*, *34*(1), 43–47.
杨强, 叶宝娟, 温忠麟. (2014b). 用 SPSS 软件计算单维测验的合成信度. *中国临床心理学杂志*, *22*(3), 496–498.
叶宝娟, 温忠麟. (2011). 单维测验合成信度三种区间估计的比较. *心理学报*, *43*(4), 453–461.
叶宝娟, 温忠麟. (2012a). 测验同质性系数及其区间估计. *心理学报*, *44*(12), 1687–1694.
叶宝娟, 温忠麟. (2012b). 用 Delta 法估计多维测验合成信度的置信区间. *心理科学*, *35*(5), 1213–1217.
叶宝娟, 温忠麟. (2013a). 有中介的调节模型检验方法: 甄别和整合. *心理学报*, *45*(9), 1050–1060.
叶宝娟, 温忠麟. (2013b). α 系数的区间估计方法比较. *心理科学*, *36*(1), 215–222.
叶宝娟, 温忠麟. (2013c). 两水平研究中单维测验信度的估计. *心理科学*, *36*(3), 728–733.
叶宝娟, 温忠麟, 陈启山. (2012). 追踪研究中测验信度的估计. *心理科学进展*, *20*(3), 467–474.
叶宝娟, 温忠麟, 胡竹菁. (2013). 单维测验合成信度元分析. *心理科学*, *36*(6), 1464–1469.
叶宝娟, 杨强. (2015). 用 Delta 法估计误差相关测验合成信度的置信区间: 以 FAD 为例. *心理学探新*, *35*(3), 251–256.
叶素静, 唐文清, 张敏强, 曹魏聪. (2014). 追踪研究中缺失数据处理方法及应用现状分析. *心理科学进展*, *22*(12), 1985–1994.
尹奎, 彭坚, 张君. (2020). 潜在剖面分析在组织行为领域中的应用. *心理科学进展*, *28*(7), 1056–1070.
余嘉元. (2005). 基于人工神经网络的一种效度凭证求取方法. *心理学报*, *37*(4), 555–560.
余嘉元. (2007). 等级效标分数的概率神经网络预测方法研究. *心理科学*, *30*(3), 666–667.
余嘉元. (2008). 粗糙集和神经网络在心理测量中的应用. *心理学报*, *40*(8), 939–946.
俞宗火, 戴海崎, 唐小娟. (2006). 全息项目因素分析在心理学研究中的应用. *心理与行为研究*, *4*(4), 306–311.
张宏宇, 许燕, 柳恒超. (2007). 社会关系模型(SRM)——个体差异研究的新策略. *心理科学进展*, *15*(6), 968–973.
张洁婷, 焦璨, 张敏强. (2010). 潜在类别分析技术在心理学研究中的应用. *心理科学进展*, *18*(12), 1991–1998.
张洁婷, 张敏强, 耿爽. (2012). 心理学研究中潜变量空间的特性探讨. *心理学探新*, *32*(5), 404–409.
张洁婷, 张敏强, 焦璨. (2019). 纳入式分类分析法在潜在

剖面模型的后续多元回归中的应用. *心理学探新*, *39*(1), 40–46.
张洁婷, 张敏强, 黎光明. (2017). 潜在剖面模型的后续分析—比较分类分析法改进后的偏差. *心理学探新*, *37*(5), 434–440.
张沥今, 陆嘉琦, 魏夏琰, 潘俊豪. (2019). 贝叶斯结构方程模型及其研究现状. *心理科学进展*, *27*(11), 1812–1825.
张沥今, 魏夏琰, 陆嘉琦, 潘俊豪. (2020). Lasso 回归: 从解释到预测. *心理科学进展*, *28*(10), 1777–1788.
张杉杉, 陈楠, 刘红云. (2017). LGM 模型中缺失数据处理方法的比较: ML 方法与 Diggle-Kenward 选择模型. *心理学报*, *49*(5), 699–710.
张珊珊, 张建新, 周明洁. (2012). 二次响应面回归方法及其在个体-环境匹配研究中的使用. *心理科学进展*, *20*(6), 825–833.
张缨斌, 王烨晖. (2019). 反应风格的测量与统计控制. *心理科学*, *42*(3), 747–754.
赵必华, 顾海根. (2010). 我国近年来心理学研究中 SEM 方法文献分析. *心理科学*, *33*(2), 403–405.
赵广平. (2015). 字母混淆矩阵的多维尺度分析. *心理学探新*, *35*(3), 284—288.
赵礼, 王晖. (2019). 统计检验力的分析流程与多层模型示例. *心理技术与应用*, *7*(5), 276–283.
郑昊敏, 温忠麟, 吴艳. (2011). 心理学常用效应量的选用与分析. *心理科学进展*, *19*(12), 1868–1878.
郑文智, 吴文毅. (2014). 结构方程模型拟合评鉴: 整体拟合、内部拟合与复核效度检验. *心理学探新*, *34*(1), 57–61.
郑显亮, 顾海根, 赵必华. (2011). 二阶因素模型的测量等价性检验——以大学生网络利他行为量表为例. *心理科学*, *34*(5), 1195–1200.
钟建军, Zoltan Dienese, 陈中永. (2017). 心理研究引入贝叶斯统计推断的必要性、应用思路与领域. *心理科学*, *40*(6), 1477–1482.
仲晓波. (2010a). 零假设检验和元分析之间的逻辑连贯性. *心理科学*, *33*(6), 1477–1480.
仲晓波. (2010b). 心理学研究中应该怎样报告实验的结果? *心理学探新*, *30*(5), 62–65.
仲晓波. (2015). 心理学实验的可重复性. *心理科学*, *38*(4), 807–812.
仲晓波. (2016). 关于假设检验的争议: 问题的澄清与解决. *心理科学进展*, *24*(10), 1670–1676.
仲晓波, 黄希尧, 万荣根. (2008). 心理学中对假设检验一些批评的分析. *心理科学*, *31*(4), 1010–1013.
周浩, 龙立荣. (2004). 共同方法偏差的统计检验与控制方法. *心理科学进展*, *12*(6), 942–950.
朱海腾. (2020). 多层次研究的数据聚合适当性检验: 文献评价与关键问题试解. *心理科学进展*, *28*(8), 1392–1408.
朱海腾, 李川云. (2019). 共同方法变异是"致命瘟疫"吗? ——论争、新知与应对. *心理科学进展*, *27*(4), 587–599.

A review of research on psychological statistics in China's mainland from 2001 to 2020

WEN Zhonglin[1], FANG Jie[2], SHEN Jiaqi[1], TAN Yitian[1], LI Dingxin[1], MA Yiming[1]
([1] *Center for Studies of Psychological Application & School of Psychology, South China Normal University, Guangzhou* 510631, *China*) ([2] *School of Humanities and Communication, Guangdong University of Finance & Economics, Guangzhou* 510320, *China*)

Abstract: A total of 213 articles on psychological statistical methods have been published in 11 journals of psychology in China's Mainland from 2001 to 2020. The research scope mainly involves the following 10 areas (ranking by the number of papers): structural equation model, test reliability, mediation effect, effect size and testing power, longitudinal study, moderation effect, exploratory factor analysis, latent class analysis, common method bias and hierarchical linear model (some articles involve more than one area). We have made a brief review and commented on each area of research. The results show that the width and depth of the research on psychological statistical methods in China are increasing, and the research hotspots are developing together in mutual integration. However, it is review articles, rather than original research articles, that account for a large proportion of the publications, and the whole research team of psychological statistical methods is not large enough.

Key words: psychological statistics, bibliometric analysis, content analysis